Addictive Behaviors Reports 8 (2018) 176–184

Contents lists available at ScienceDirect

Addictive Behaviors Reports

journal homepage: www.elsevier.com/locate/abrep

Psychometric assessment of the Internet Gaming Disorder diagnostic criteria: An Item Response Theory study

Bruno Schivinski[a,*], Magdalena Brzozowska-Woś[b], Erin M. Buchanan[c], Mark D. Griffiths[d], Halley M. Pontes[d]

[a] Birkbeck, University of London, Department of Management, Malet Street, Bloomsbury, WC1E 7HX London, United Kingdom of Great Britain and Northern Ireland
[b] Gdansk University of Technology, Department of Marketing, Ul. Narutowicza 11/12, Gdansk 80-233, Poland
[c] Missouri State University, Department of Psychology, 901 S. National Ave, Springfield, MO 65897, USA
[d] Nottingham Trent University, International Gaming Research Unit, Psychology Department, 50 Shakespeare Street, Nottingham NG1 4QF, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:
Internet Gaming Disorder
Video games
DSM-5
IGDS9-SF
Behavioral addictions

ABSTRACT

Internet Gaming Disorder (IGD) has been recognized by the American Psychiatric Association (APA) as a tentative disorder in the latest fifth revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). In order to advance research on IGD, the APA has suggested that further research on the nine IGD criteria to investigate its clinical and empirical feasibility is necessary. The aim of the present study was to develop the Polish the Internet Gaming Disorder Scale–Short-Form (IGDS9-SF) and scrutinize the nine IGD criteria empirically. To achieve this, the newly developed IGDS9-SF was examined using a wide range of psychometric methods, including a polytomous Item Response Theory (IRT) analysis to evaluate the measurement performance of the nine IGD criteria. A sample of 3377 gamers (82.7% male, mean age 20 years, SD = 4.3 years) was recruited online for the present study. Overall, the findings obtained confirmed that suitability of the Polish IGDS9-SF to assess IGD amongst Polish gamers given the adequate levels of validity and reliability found. The IRT analysis revealed that the IGDS9-SF is a suitable tool to measure IGD levels above the average; however, criteria "continuation" (item 6), "deception" (item 7), and "escape" (item 8) presented with poor fit. Taken together, these results suggest that some of the diagnostic criteria may present with a different clinical weighting towards final diagnosis of IGD. The implications of these findings are further discussed.

1. Introduction

Internet Gaming Disorder (IGD) is defined by a persistent and recurrent involvement with videogames, often leading to significant impairments of daily, work and/or educational activities and has been suggested by the American Psychiatric Association (APA) as a tentative psychiatric disorder requiring further study (*Diagnostic and Statistical Manual of Mental Disorders* [DSM-5] (APA, 2013). According to the DSM-5 (APA, 2013), IGD is indicated by endorsement of at least five core symptoms (out of nine) over a period of 12 months. More specifically, the diagnostic criteria of IGD include the following nine clinical symptoms: (1) preoccupation with videogames (i.e., "*preoccupation*"); (2) experience of unpleasant symptoms when videogaming is taken away (i.e., "*withdrawal*"); (3) the need to spend increasing amounts of time engaged in videogames (i.e., "*tolerance*"); (4) unsuccessful attempts to control participation in videogames (i.e., "*loss of control*"); (5) loss of interest in previous hobbies and entertainment as a result of, and with the exception of, videogames (i.e., "*give up other activities*"); (6) continued excessive use of videogames despite knowledge of psychosocial problems (i.e., "*continuation*"); (7) deceiving family members, therapists, or others regarding the amount of videogaming (i.e., "*deception*"); (8) use of videogames to escape or relieve negative moods (i.e., "*escape*"); and (9) jeopardizing or losing a significant relationship, job, or education or career opportunity because of participation in videogames (i.e., "*negative consequences*").

Epidemiological research investigating the extension of problems caused by disordered gaming across a number of countries found a relatively low prevalence rate of IGD. Based on the findings reported by robust studies with large and representative samples, IGD was estimated to affect about 5.7% of German adolescents and young adults aged between 12 and 25 years (Wartberg, Kriston, & Thomasius, 2017). In Korea, IGD has been found to affect about 4.0% of adults aged from 18 to 39 years (Park, Jeon, Son, Kim, & Hong, 2017) and 5.9% of adolescents aged between 13 and 15 years (Yu & Cho, 2016). Another

* Corresponding author.
E-mail addresses: bruno.schivinski@gmail.com, b.schivinski@bbk.ac.uk (B. Schivinski).

https://doi.org/10.1016/j.abrep.2018.06.004
Received 22 March 2018; Received in revised form 4 May 2018; Accepted 20 June 2018
Available online 30 June 2018

recent study reported a prevalence rate of IGD about 2.1% amongst Slovenian adolescents 13 years old (Pontes, Macur, & Griffiths, 2016b), and in Norway IGD is estimated to affect around 1.4% of individuals aged between 16 and 74 years old (Wittek et al., 2016). Although IGD appears to affect a minority of individuals, its investigation is still warranted given the wide detrimental outcomes generally associated with disordered gaming.

At the psychosocial and behavioral level, a wide range of negative outcomes associated with IGD have been consistently reported by a large body of studies. These include lower social support and health-related quality of life (Wartberg, Kriston, & Kammerl, 2017), antisocial behavior, anger control problems, emotional distress, decreased self-esteem, and attention deficit hyperactivity disorder (Wartberg et al., 2017), social phobia (Sioni, Burleson, & Bekerian, 2017), decreased satisfaction with life and low self-efficacy (Festl, Scharkow, & Quandt, 2013), poor academic performance (Brunborg, Mentzoni, & Frøyland, 2014), denial coping strategy (Schneider, King, & Delfabbro, 2017), nicotine use disorder, depression and anxiety (Park et al., 2017), interpersonal problems (Arcelus et al., 2017), and impaired psychological wellbeing (Lim et al., 2016).

Despite the relatively high amount of empirical and clinical studies conducted on IGD more recently (e.g., Frölich et al., 2016; Leménager et al., 2016; Li et al., 2017; Pontes, Király, Demetrovics, & Griffiths, 2014; Pontes, Kuss, & Griffiths, 2015), there is still a substantial amount of inconsistencies in its nosological features due to limited data regarding the clinical course and etiology of this particular condition (Kuss, Griffiths, & Pontes, 2017a, 2017b). Therefore, it is unlikely that the APA will formally recognize IGD until future research has: (i) identified its defining features, (ii) obtained cross-cultural data on reliability and validity of its diagnostic criteria, (iii) determined prevalence rates in representative epidemiological samples in countries around the world, and (iv) evaluated its natural history and examine its associated biological features (Petry & O'Brien, 2013).

In addition to these issues, the limited amount of studies conducted to date examining the psychometric properties of each of the nine diagnostic criteria of IGD produced mixed findings that warrant further investigation. For instance, Rehbein, Kliem, Baier, Mößle, and Petry (2015) reported that the criteria "*give up other activities*", "*tolerance*", and "*withdrawal*" were the most useful and informative criteria. Moreover, Lemmens, Valkenburg, and Gentile (2015) reported that "*escape*" did not add diagnostic accuracy as this criterion lacked specificity. A more recent study conducted by Király et al. (2017) concluded that "*continuation*", "*preoccupation*", "*negative consequences*", and "*escape*" were more associated with lower levels of IGD while "*tolerance*", "*loss of control*", "*give up other activities*", and "*deception*" where more informative at higher levels of IGD. There may be a few potential explanations for these distinct findings that merit consideration. On the one hand, these studies relied on different assessment tools to evaluate IGD. On the other hand, the samples recruited for these studies differed systematically in their basic demographic features in addition to the different sampling techniques utilized across these studies (e.g., non-probability sampling and probability sampling).

As a way to mitigate these methodological issues, adopting well-established and validated psychometric tools to assess IGD is necessary (see Griffiths, King, & Demetrovics, 2014; Petry & O'Brien, 2013; Pontes & Griffiths, 2014; Pontes, Kuss, & Griffiths, 2017). For this reason, the Internet Gaming Disorder Scale–Short-Form (IGDS9-SF; Pontes & Griffiths, 2015) was developed based upon the nine DSM-5 criteria aforementioned (APA, 2013). Since its initial development, a number of studies have employed this tool to assess IGD across a wide range of cultural contexts and samples, such as Portuguese (Pontes & Griffiths, 2016), Slovenian (Pontes et al., 2016b), Italian (Monacis, De Palo, Griffiths, & Sinatra, 2016), and Persian (Wu, Lai, Yu, Lau, & Lei, 2017; Wu, Lin, et al., 2017). Overall, these studies reported consistent findings supporting a one-factor model for the IGDS9-SF alongside its suitability to assess IGD in different cultural contexts (Stavropoulos et al., 2017).

The IGDS9-SF has been used extensively internationally and most previous studies investigating the psychometric properties of this tool were exclusively based on the use of Classical Test Theory (CTT) by conducting exploratory factor analysis (e.g., Pontes & Griffiths, 2015), confirmatory factor analysis (CFA) (e.g., Pontes, Macur, & Griffiths, 2016a), and multiple linear regression (e.g., Monacis, de Palo, Griffiths, & Sinatra, 2017). There is, however, an emerging body of literature on the IGDS9-SF that have attempted to provide information regarding its psychometric properties using an Item Response Theory (IRT) approach (e.g., Gomez, Stavropoulos, Beard, & Pontes, 2018; Király et al., 2017). Despite this emerging IRT-based literature, the field still lacks further psychometric information on the IGDS9-SF at the item level using for example a robust IRT approach since CTT-based approaches are not able to shed light on this issue.

1.1. Item Response Theory

The application of IRT models help determine the precision of psychological scales by identifying two key parameters of an item when using a 2-parameter logistic model (2-PL). The first of these parameters is the *location parameter* (β). When answer choices are considered correct/incorrect, the location parameter is often interpreted as item difficulty. However, when using multiple response formats, item location parameters indicate the threshold or level of latent trait between answer choices, thus revealing the level of latent variable wherein someone might choose that response option. The second parameter is the *discrimination or precision parameter* (α), which reflects the degree to which an item discriminates individuals across the latent trait, where high values indicate steeper discrimination slopes. Contrary to CTT-based approaches, IRT models assume that the items of a scale are not equally informative across the latent trait range. In fact, in IRT models, one item can provide more or less information than another item according to both parameters (e.g., α and β). Since the coefficient alpha of two scales might be the same under CTT assumptions, this psychometric approach does not help informing how precise an instrument is at different levels of the latent trait theta (θ) (Ayesrst & Bagby, 2011).

In order to calculate an individual's score on a particular latent trait, IRT models need first to identify the item location thresholds (β) and discrimination parameters (α) for all items of a scale. This calibration is achieved by using the responses of the whole sample to estimate each item parameters. Moreover, the location thresholds (β) and discrimination (α) parameters can then be used to generate a test information function (TIF), which provides an estimate of the precision of the entire scale across the trait being measured. Note that the peak in the TIF occurs when measurement precision is greatest as the most information in the measurement is identified. If equally precise measurement is desirable across a latent trait range (θ), then a relatively flat curve is desirable as it indicates that the items of a scale are highly discriminatory across a range of the trait symptoms or severity being measured (Embretson & Reise, 2000). The latent trait examined in IRT models (θ) is similar to a *z*-score and is assumed to have a mean of 0 and standard deviation of 1 (Ayesrst & Bagby, 2011). Based on this standardization, clinical research has suggested that the desired TIF would provide information across all levels of a specific trait and might include a range from 2 SD below the mean to 2 SD above the mean (Sibley, 2012).

1.2. The current study

To the best of the authors' knowledge, no previous study using polytomous IRT on the nine diagnostic criteria for IGD has been conducted. Since polytomous IRT provides additional psychometric information that is significantly distinct from the psychometric information derived CTT-based research (see Hambleton & Jones, 1993 for a discussion), such investigation may have the potential to yield fruitful insights that may help inform official medical bodies (e.g., APA) about the specific diagnostic properties the nine IGD criteria (Petry & O'Brien, 2013; Petry, Rehbein, Ko, & O'Brien, 2015).

In light of the aforementioned, the aim of the present study was twofold. First, to translate and validate the Polish version of the IGDS9-SF and provide further cross-cultural information about the scale. Second, to determine the extent to which the IGDS9-SF can reliably differentiate between individuals at different levels of the IGD latent trait and provide further psychometric evidence using polytomous IRT. It is envisaged by the present authors that this study will contribute to the broader international discussion on the usefulness of the nine IGD criteria for assessing disordered gaming (see Kaptsis, King, Delfabbro, & Gradisar, 2016a, 2016b; King & Delfabbro, 2016; Starcevic, 2017).

2. Method

2.1. Participants and procedures

In the present study, we aimed at recruiting a large and heterogeneous sample of online gamers from Poland. In order to achieve this goal, administrators from the three most popular gaming forums in Poland (i.e., www.gry-online.pl, www.gamesboard.pl, and www.gamesfanatic.pl) were individually contacted and invited to collaborate with the research team by assisting the researchers in the recruitment process of the sample. After obtaining permission from all three forum administrators, data collection was carried out by disseminating a link to an online survey created and hosted on Qualtrics (www.qualtrics.com) containing the study's main psychometric instruments. More specifically, the link of the survey was disseminated by all three gaming forums through their email subscriptions service, threads on each online forum, and via their official social media channels (e.g., Facebook, Twitter, and YouTube). The data collection period spanned from March 23rd until May 24th, 2017.

During the data collection stage, the link of the study's survey was advertised on a weekly basis using the social media channels from all three forums aforementioned in order to recruit gamers to take part in the survey. After clicking on the survey's link, respondents were redirected to the study's questionnaire and briefed about their right to remain anonymous and confidential. Moreover, participants' eligibility was initially verified by asking if they had played videogames in the past 12 months, with those answering '*no*' to this question being removed from the analysis. As a result, a total of 110 participants (3.2%) were removed on the basis of this initial inclusion criterion. Upon completion of the recruitment process, a total of 3377 participants were successfully recruited. Overall, the sample's mean age was 20 years (SD = 4.3 years, range 12–49 years) and in terms of gender distribution, males represented 82.67% (n = 2789) of the total sample.

The study was approved by the ethics committees of the research team's institutions, and electronic consent was obtained from all participants as a requirement to partake in the present study. Furthermore, all procedures were followed in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2005.

2.2. Measures

2.2.1. Sociodemographics and gaming-related behaviors

The survey included questions that were comparable to previous similar psychometric studies (e.g., Monacis et al., 2016; Pontes & Griffiths, 2016; Stavropoulos et al., 2017) using the IGDS9-SF (Pontes & Griffiths, 2015). Thus, sociodemographic data of participants' gender, age, and relationship status were collected. Gaming-related behaviors were assessed by three questions asking participants about their average time spent gaming during the weekdays (Monday to Friday) and weekends (Saturday and Sunday), and average time spent per gaming session. Finally, the survey included a question asking if participants played any videogames from their smartphones (yes/no), if they agreed with following statement: '*I would consider myself addicted to video games.*' (answers ranged from 1 'Strongly disagree' to 5 'Strongly agree'), and if they were active gamers (yes/no).

2.2.2. Internet Gaming Disorder Scale–Short-Form (IGDS9-SF)

The nine-item IGDS9-SF (Pontes & Griffiths, 2015) is a short psychometric tool based on the nine core criteria defining IGD as suggested by the DSM-5 (American Psychiatric Association, 2013). This tool assesses the severity of IGD and its detrimental effects by examining both online and/or offline gaming activities occurring over a 12-month period, and all of its nine items are answered using a 5-point scale: 1 ('*Never*'), 2 ('*Rarely*'), 3 ('*Sometimes*'), 4 ('*Often*'), and 5 ('*Very often*'). Participants' total scores can be obtained by summing participants' responses to the nine items and can range from 9 to 45, with higher scores being indicative of a higher degree of disordered gaming. As suggested by Pontes and Griffiths (2015), in order to distinguish between disordered and non-disordered gamers, researchers can operationalize endorsement of the nine diagnostic criteria by recoding responses given by participants as 5 ('*Very often*') as endorsement of specific criterion.

In order to develop the Polish version of the IGDS9-SF, standard procedural methods and guidelines used for the process of cross-cultural adaptation of self-report measures were adopted (Beaton, Bombardier, Guillemin, & Ferraz, 2000). The forward translation process of the IGDS9-SF was carried out by two independent bilingual translators whose mother tongue was Polish, this procedure involved translating the original English version of the IGDS9-SF to Polish. Discrepancies that have emerged between the two translations were dealt by the research team members that are fluent in Polish. Following this procedure, an interim version of the Polish IGDS9-SF was generated, and then back-translated into English by two independent native English speakers that were not aware of the IGDS9-SF. In order to ensure the contents of the scale were preserved (i.e., semantic properties) in the Polish IGDS9-SF, the back-translated versions were compared with the original IGDS9-SF. Following this, the translated versions of the IGDS9-SF were then consolidated in a session carried out by the research team and all translators. Finally, the Polish version of IGDS9-SF was piloted by 52 potential test-takers (51% male, $\text{Mean}_{\text{age}} = 21.4$ years, SD = 3.5 years) who shared their perceptions and interpretations of each of the nine Polish items of the IGDS9-SF, this procedure aided the assessment of facial and content validity of the Polish IGDS9-SF (see Appendix 1).

2.3. Data management and analytic strategy

Data management involved cleaning the dataset by inspecting cases with severe missing values across the IGDS9-SF. Multiple imputation was carried out whenever missing was at random cases missing up to 2 items out of the nine IGDS9-SF items (n = 117, average replacement of 0.4%) using the package mice (Multivariate Imputation by Chained Equation Version 2.3) (Van Buuren & Groothuis-Oudshoorn, 2011) in R system for statistical computing Version 3.4.1 (https://www.r-project.org). Furthermore, a total of 155 (4.6%) cases were excluded from the analyses due to presenting severe missing values on ≥3 items of the IGDS9-SF.

As for the assessment of univariate normality, no item of the IGDS9-SF had absolute values of Skewness > 3.0 and Kurtosis > 8.0 (Kline, 2011). In order to screen for univariate outliers, a standardized composite sum score of the IGDS9-SF using all nine items was created and participants were deemed univariate outliers if they scored ± 3.29 standard deviations from the IGDS9-SF z-scores. This threshold was adopted because it includes around 99.9% of the normally distributed IGDS9-SF z-scores (Field, 2013). The data were also screened for multivariate outliers using Mahalanobis distances and the critical value for each case based on the chi-square distribution values, which resulted in no further exclusion of participants. Finally, the data were also checked for accuracy and missing values (i.e., missing two or less items on the IGDS9-SF). As a result of the aforementioned data cleaning procedures, a final sample size of 3222 (95.4%) participants was eligible for all the subsequent analyses.

2.4. Statistical analyses

The statistical analyses comprised descriptive analysis of the characteristics of the sample (i.e., frequencies and percentages); construct validity and unidimensionality assessment of the IGDS9-SF via Confirmatory Factor Analysis (CFA); reliability analysis of the IGDS9-SF using Cronbach's alpha and Composite Reliability (CR) and Factor Determinacy (FD) coefficients of internal consistency; criterion validity analysis of the IGDS9-SF by estimating the CFA model with covariates in a Multiple Indicator, Multiple Cause (MIMIC) model including the latent construct of IGD and observable measures such as participants' self-reported frequency of gameplay. The statistical analyses outlined above were conducted using IBM SPSS Version 23 (IBM Corp, 2015) and *Mplus* 7.2 (Muthén & Muthén, 2012)

Further psychometric testing was carried out through polytomous IRT in order to ascertain which items of the IGDS9-SF were moderately to highly discriminant and better suited for identifying moderate levels of gaming disorder. For the IRT analyses, local independence was examined by using the Q_3 statistic with a critical score of correlated residuals > 0.20 (Christensen, Makransky, & Horton, 2017; Yen, 1984). One residual correlation between items 3 and 9 was found to violate this assumption (0.23), indicating that these two items are somewhat dependent even after accounting for latent traits (Christensen et al., 2017; Yen, 1984). All other residual correlations were below 0.20. Monotonicity assumption was examined by plotting trace curves for each item, examining the ordering of Likert responses and item fit characteristics. The IRT analysis was performed with R system for statistical computing using the mirt package (Multidimensional Item Response Theory Version 1.25) and the default dimensional reduction expectation/maximization (EM) algorithm suggested for unidimensional models (Bock & Aitkin, 1981; Chalmers, 2012).

3. Results

3.1. Descriptive statistics

With regards to the age of participants, about 21% (n = 677) were 12–16 years old, 69.2% (n = 2230) were 17–25 years old, 8.5% (n = 274) were 26–37 years old, and the remainder of participants were 38–46 years old (0.6% n = 20). As for participants' relationship status, 70.5% (n = 2268) reported not being in a romantic relationship.

Further descriptive analysis indicated that the majority of sample (65.3% n = 2203) had been playing videogames for an average of 8 years (SD = 2.9 years). The average time spent playing videogames was 7.5 h for weekdays (SD = 6.73 h) and 7.17 h for weekends (SD = 5.31 h), further reflecting an average of about 2.79 h per gaming session (SD = 2.12 h). Finally, about 22.3% (n = 717) of the sample reported using their smartphones to play videogames, and a total of 14.4% (n = 464) either '*agreed*' or '*strongly agreed*' to the following statement: "*I would consider myself addicted to videogames*". Finally, the vast majority of the sample considered themselves to be gamers (76.4% n = 2462).

3.2. Construct validity and unidimensionality

Construct validity and unidimensionality of the IGDS9-SF was investigated by performing a CFA with robust maximum likelihood estimation method (MLR) on its nine items to test the one-factor solution as previously reported (Monacis et al., 2016; Pontes et al., 2016b; Pontes & Griffiths, 2015; Pontes & Griffiths, 2016; Wu, Lai, et al., 2017; Wu, Lin, et al., 2017). Conventional fit indices and thresholds were used to examine the goodness of fit of the model under analysis: χ^2/d.f. [1;4]; Root Mean Square Error of Approximation (RMSEA) [0.05;0.08], RMSEA 90% Confidence Interval (CI) with its lower limit close to 0 and the upper limit below 0.08; probability level value of the test of close fit (Cfit) > 0.05; Standardized Root Mean Square Residual (SRMR) [0.05;0.08]; Comparative Fit Index (CFI); and Tucker-Lewis Fit Index (TLI) [0.90;0.95] (Bentler, 1990; Bentler & Bonnet, 1980; Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999). The CFA yielded the following results: $\chi^2_{(27)}$ = 188.753; χ^2/df = 6.99; CFI = 0.968; TLI = 0.958; RMSEA = 0.043 [90% CI: 0.037–0.049; Cfit = 0.98]; SRMR = 0.025. Overall, the results of the CFA produced acceptable standardized item loads (i.e., $\lambda_{ij} \geq 0.50$, $p < 0.0001$), with the exception of criterion '*escape*' (λ_{item8} = 0.47, $p < 0.0001$) (see Fig. 1). However, this result was not deemed problematic given that standardized loadings between 0.45 and 0.54 are considered '*fair*' in social science (Tabachnick & Fidell, 2013).

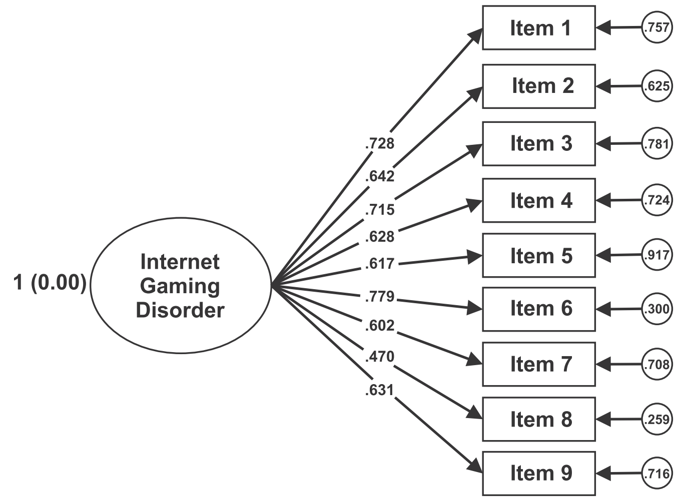

Fig. 1. Graphical summary of the confirmatory factor analysis results obtained from the nine items of the IGDS9-SF (N = 3,222).

3.3. Criterion-related validity

Criterion validation of the IGDS9-SF was investigated with a MIMIC model whereby IGD was predicted by the following gaming-related behaviors: average time spent playing videogames during the weekdays and weekends, and average time spent per gaming session. These gaming-related behaviors were selected as the three main external criteria. The criteria choice should be based on variables that are reliable indicators of the trait the instrument intends to measure (Rubin & Babbie, 2009). Several studies have shown that IGD is commonly associated with greater frequency of gameplay (Fuster, Carbonell, Pontes, & Griffiths, 2016; Lemmens, Valkenburg, & Peter, 2009; Pápay et al., 2013; Pontes & Griffiths, 2015; Schneider et al., 2017).

In line with the previous findings, the results demonstrated that IGD was positively influenced by time-spent gaming during the weekdays (β = 0.08, p = 0.001), weekends (β = 0.36, p = 0.001), and average time spent per gaming session (β = 0.09, p = 0.001). Furthermore, the MIMIC model presented with excellent fit to the data ($\chi^2_{(51)}$ = 275.323; χ^2/df = 5.39; CFI = 0.963; TLI = 0.954; RMSEA = 0.037 [90% CI: 0.033–0.041; Cfit = 1]; SRMR = 0.025) (see Fig. 2), lending further empirical support to the measure's criterion-related validity.

3.4. Reliability analysis

The internal consistency of the scale was evaluated using different coefficients of reliability. More specifically, reliability level estimated by the Cronbach's alpha was excellent (α = 0.82) while the CR

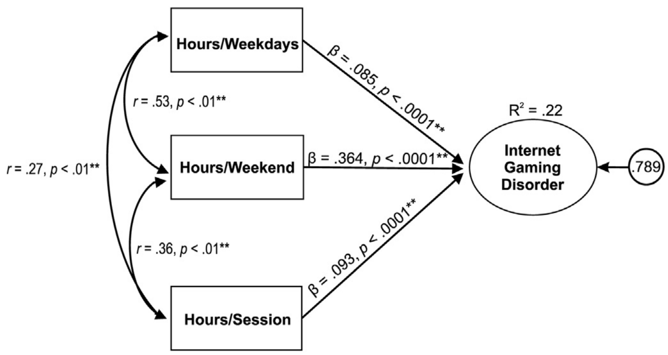

Fig. 2. Summary of the criterion-related validity analysis using MIMIC model predicting overall Internet Gaming Disorder scores by average time spent playing videogames during weekdays and weekends, and average time spent per gaming session (N = 3,222).

coefficient was 0.87, which is well above the recommended threshold of 0.70 (Fornell & Larcker, 1981; Hair, Black, Babin, & Anderson, 2010). Finally, reliability as measured by the FD coefficient was 0.91, which is above the desired threshold of 0.80 (Muthén & Muthén, 2012). Taken together, these results suggest that the Polish version of the IGDS9-SF presents with excellent internal consistency levels given the high coefficients obtained, a finding that echoes the findings of similar studies using the IGDS9-SF in other countries (e.g., Monacis et al., 2016; Pontes et al., 2016b; Pontes & Griffiths, 2015; Pontes & Griffiths, 2016; Wu, Lai, et al., 2017; Wu, Lin, et al., 2017).

3.5. IRT Analysis

In order to further investigate the diagnostic properties of the IGDS9-SF from a psychometric perspective, a graded response model (Samejima, 1997) was analyzed and compared to a generalized partial credit model (Muraki, 1992). The 2-PL model is calculated by solving for α and β:

$$P(\mu = 1|\ \theta\Big) = \frac{e^{\alpha(\theta-\beta)}}{1 + e^{\alpha(\theta-\beta)}}$$

that creates an item characteristic curve denoting the slope α and location β for each item. This formula can be modified to examine multiple response formats (polytomous models) by calculating the thresholds between categories, rather than examining one incorrect/correct discrimination. The graded response model (Samejima, 1997) is the simplest model, where each response option threshold is compared to all response options above that level. The threshold indicates the location on θ at which individuals would be equally likely to indicate comparison response options. Therefore, β_1 is a 2-PL using answer choice 1 versus all other answer choices (i.e., 2, 3, 4, and 5), while β_2 indicates the threshold between answer choice 2 and 3, 4, and 5. In comparison, a generalized partial credit model (Muraki, 1992) does not assume that response options are ordered, and instead calculates β values by comparing adjacent response options. Thus, β_1 would indicate the location of answer choice 1 to 2, while β_2 would indicate the location for answer choice 2 to 3, respectively. To determine which model approximated the data, the Akaike Information Criterion (AIC) was used to compare a graded response and generalized partial credit model. Models with lower AIC values are desirable, as they indicate a closer fit to a true model (Burnham & Anderson, 2002). The results indicated that the graded response model showed better fit to the data ($AIC_{graded\ model} = 70{,}167.85$; $AIC_{graded\ partial\ credit\ model} = 70{,}354.38$). Table 1 includes parameter estimates for the model, as well as item fit statistics.

A visual inspection of the trace curves for each item indicated that the items of the IGDS9-SF appear monotonic (see Fig. 3), wherein Likert choices of the scale are ordered. Item location threshold values (β) can be found in Table 2. In addition to ordering, these curves were examined to determine if each answer choice showed an area of latent trait that they were most probable. As a result, criteria "*give up other activities*" (item 5) and "*escape*" (item 8) showed that choice 2 '*rarely*' was depressed in relation to other item choices, thus, portraying that individuals were unlikely to select *'rarely'* in comparison to other answer choices. Analysis of the discrimination levels (α) for each item was overall high (i.e., < 1.0) with the lowest level criterion "*escape*" ($\alpha_{item\ 8} = 0.88$). Criteria '*preoccupation*' ($\alpha_{item\ 1} = 1.72$) and '*withdrawal*' ($\alpha_{item\ 2} = 1.71$) were the strongest discrimination items, followed by '*tolerance*' ($\alpha_{item\ 3} = 1.60$), '*continuation*' ($\alpha_{item\ 6} = 1.59$), '*negative consequences*' ($\alpha_{item\ 9} = 1.56$), '*loss of control*' ($\alpha_{item\ 4} = 1.53$), '*deception*' ($\alpha_{item\ 7} = 1.48$), *and* '*give up other activities*' ($\alpha_{item\ 5} = 1.38$). Item fit was calculated using the generalized S-χ^2 statistic (Kang & Chen, 2008; Orlando & Thissen, 2000), which calculates the sum of the differences between observed and expected proportion of responses for each test score. Significant scores indicate potential misfit to the model; however, these statistics are also sensitive to large sample sizes given the N multiplier in their formula. Item fit was acceptable for most items, with the poorest fit for criteria "*continuation*" (S-$\chi^2_{item\ 6} = 132.65$, $p = 0.001$), "*deception*" (S-$\chi^2_{item\ 7} = 125.26$, $p = 0.005$), and "*escape*" (S-$\chi^2_{item\ 8} = 177.50$, $p \leq 0.001$).

The test information curve [I(θ)] is calculated by summing information provided by each item across θ:

$$I(\theta) = \sum_i \alpha_i^2 P_j(\theta, \beta_{ij}, \alpha_i) Q_j(\theta, \beta_{ij}, \alpha_i)$$

P_j (θ, β, α) indicates the likelihood of selecting a response choice (j) at a given latent trait level for each item (i), while Q_j (θ, β, α) calculates the likelihood of *not* selecting that response choice [i.e., 1 - P_j (θ, β, α)]. Fig. 4 contains the test information curve estimated and indicates that overall the IGDS9-SF is useful at measuring IGD levels above the mean (i.e., the curve peaks at z-scored latent traits of 0–2), which is especially necessary for a clinical diagnostic tool. Information below the mean is present, covering a range close to the desired −2 to +2 for diagnosticity. These results were replicated when the dataset was examined for participants with complete data (i.e., no missing data imputation) and for smaller subsamples of the data (i.e., randomly sampling one third of the dataset).

4. Discussion

The present study aimed at determining the extent to which the Polish version of the IGDS9-SF can be a valid and reliable psychometric tool to assess IGD in Polish speaking samples, and evaluating the scale's ability to differentiate between individuals at different levels of the IGD latent trait via a polytomous IRT analysis. Overall, the results from the psychometric analyses conducted provided support for the scale's validity at the construct, criterion, and factorial validity level. Similarly to what has been found in previous studies using the IGDS9-SF in international samples, a one-factor solution was found in the present study. This finding concurs with a large body of recent studies using the IGDS9-SF in Portugal (Pontes & Griffiths, 2016), Slovenia (Pontes et al., 2016b), Italy (Monacis et al., 2016), Iran (Wu, Lai, et al., 2017; Wu, Lin, et al., 2017), Australia, India, United States of America, and the United Kingdom (Stavropoulos et al., 2017).

Further psychometric analysis indicated satisfactory levels of overall construct and criterion-related validity according to the chosen external criteria. Although the magnitude of the associations between the three gaming-related behaviors (i.e., average time spent playing videogames during the weekdays and weekends, and average time spent per gaming session) and the levels of IGD were not exceedingly high, they were in the expected direction and thus echo previous findings. Wu, Lai, et al. (2017) and Wu, Lin, et al., 2017assessed IGD symptoms with the IGDS9-SF in a representative sample of 2363 Iranian

Table 1
Overall descriptive statistics across all items of the Polish IGDS9-SF (N = 3222).

Item	λ_{ij}	*t*	Mean (SD)	Sk	Ku	R^2
1. Do you feel preoccupied with your gaming behavior? (Example: Do you think gaming has become the dominant activity in your daily life?)	0.73	45.11	2.12 (1.13)	0.71	−0.42	0.41
2. Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?	0.64	41.50	1.84 (1.01)	1.15	0.74	0.39
3. Do you feel the need to spend increasing amount of time engaged gaming in order to achieve satisfaction or pleasure?	0.72	44.09	2.17 (1.13)	0.73	−0.27	0.39
4. Do you systematically fail when trying to control or cease your gaming activity?	0.63	37.29	1.91 (1.05)	1.14	0.70	0.35
5. Have you lost interests in previous hobbies and other entertainment activities as a result of your engagement with the game?	0.62	32.29	1.71 (1.13)	1.51	1.19	0.29
6. Have you continued your gaming activity despite knowing it was causing problems between you and other people?	0.78	41.25	2.15 (1.27)	0.82	−0.49	0.37
7. Have you deceived any of your family members, therapists or others because the amount of your gaming activity?	0.60	36.33	1.89 (1.03)	0.96	0.09	0.33
8. Do you play in order to temporarily escape or relieve a negative mood (e.g., helplessness, guilt, anxiety)?	0.47	21.49	2.88 (1.21)	−0.08	−0.91	0.14
9. Have you jeopardized or lost an important relationship, job or an educational or career opportunity because of your gaming activity?	0.63	37.23	1.84 (1.05)	1.12	0.40	0.35

Abbreviations: IGDS9-SF: The Internet Gaming Disorder Scale–Short-Form; λ_{ij}: standardized factor loading; *t*: *t*-test statistic; SD: standard deviation; Sk: skewness; Ku: kurtosis; R^2: R-squared.

adolescents and found a positive association between levels of IGD and weekly hours spent on gaming. In a similar vein, a recent study conducted by Hawi and Samaha (2017) in a sample of 375 Lebanese students also found IGD levels (measured with the Internet Gaming Disorder Test [IGD-20 Test]; Pontes et al., 2014) to be positively associated with greater time spent playing videogames. In terms of reliability, the analyses conducted provided excellent reliability coefficients at different levels, further supporting the findings of previous similar studies using the IGDS9-SF. Taken together, these findings lend further support to the validity and reliability of the IGDS9-SF and illustrate its suitability to assess IGD in Polish speaking samples.

Another aim of the present study was to contribute to the broader discussion about the validity of each of the nine IGD criteria given the mixed findings previously reported. Overall, the polytomous IRT analysis indicated that the IGDS9-SF is a suitable psychometric tool to assess IGD as the scale was able to measure the high end of the latent trait IGD with the most information possible. This finding is important given that IGD is a tentative psychiatric condition that needs to be clinically assessed effectively. Although the IGDS9-SF appears to be a suitable measure for assessing IGD levels above the mean (e.g., clinical samples with severe symptoms), which is particularly necessary for a clinical diagnostic tool, some criteria appeared to be psychometrically problematic.

More specifically, criteria 7 (i.e., "*deception*"; IGDS9-SF item 7) and 8 (i.e., "*escape*"; IGDS9-SF item 8) presented with the poorest fit in

Table 2
Item statistics for the graded response model across all items of the Polish IGDS9-SF (N = 3222).

Item	α	β_1	β_2	β_3	β_4	$S\text{-}\chi^2$	df	*p*
Item 1	1.72	−0.37	0.54	1.56	2.65	77.34	83	0.655
Item 2	1.71	−0.04	1.02	2.01	2.92	97.36	83	0.134
Item 3	1.60	−0.53	0.55	1.61	2.57	88.98	84	0.334
Item 4	1.53	−0.16	1.01	2.04	2.84	101.69	88	0.151
Item 5	1.38	0.55	1.24	1.96	2.85	96.28	91	0.333
Item 6	1.59	−0.21	0.61	1.38	2.22	132.65	84	0.001
Item 7	1.48	−0.09	0.91	2.09	3.39	125.26	87	0.005
Item 8	0.88	−1.94	−0.82	0.90	2.94	177.50	85	< 0.001
Item 9	1.56	0.06	0.98	1.99	3.12	112.32	86	0.030

Note: Chi-square statistics are fit using $S\text{-}\chi^2$: generalized chi-square statistic. (Kang & Chen, 2008; Orlando & Thissen, 2000.)
Abbreviations: α: discrimination parameter; β: difficulty parameter; χ^2: Chi-square statistic.

comparison to the remaining criteria. In one of the few clinical samples where the nine IGD criteria were assessed, it was found that criteria 7 (i.e., "*deception*") and 8 (i.e., "*escape*") presented with the poorest diagnostic accuracy (Ko et al., 2014). Thus, the present findings substantiate the results reported by Ko et al. (2014). There are a few potential explanations for this finding at the theoretical level. First, criterion 7 (i.e., "*deception*") may understood to be socially dependent,

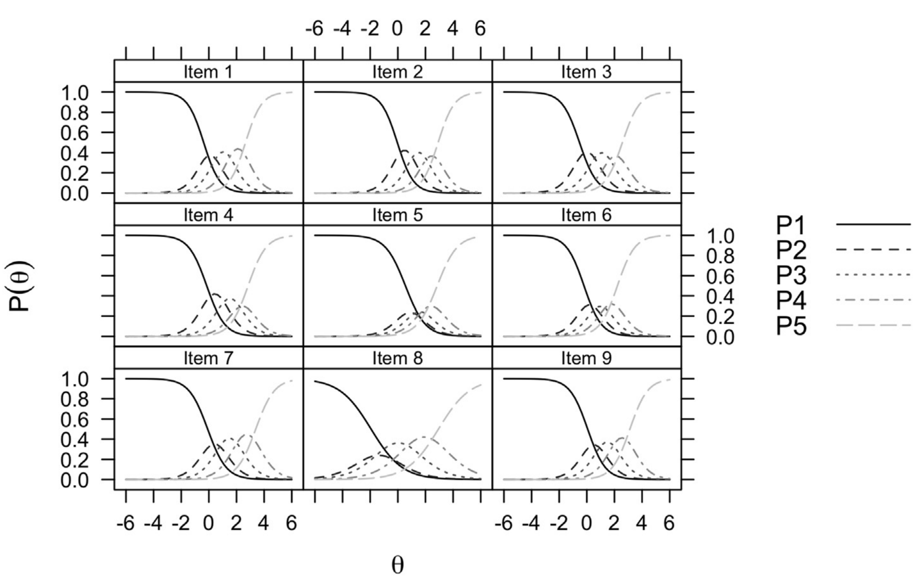

Fig. 3. Trace curves for the items of the IGDS9-SF.

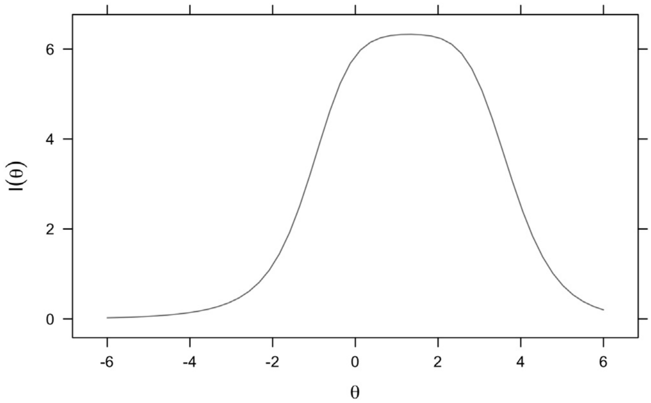

Fig. 4. Test information curve for the overall IGDS9-SF.

hence it may not be entirely appropriate to measure IGD-related symptoms in gamers without a therapist and/or significant others living with them. Second, with regards to criterion 8 (i.e., "*escape*"), a large body of empirical studies reported that "*escape*" is one of the main motives for videogame playing (e.g., Ballabio et al., 2017; Kahn et al., 2015; Yee, 2006), and further studies (e.g., Bijvank, Konijn, & Bushman, 2012; Hagström & Kaldo, 2014; Király et al., 2015; Männikkö, Billieux, Nordström, Koivisto, & Kääriäinen, 2017; Wu, Lai, et al., 2017) found "*escape*" to be the motivational factor mostly associated with disordered gaming. These studies illustrate the intricacies between disordered gaming and "*escape*" as it is not clear the role of this variable in the etiology of IGD. The role of "*escape*" in IGD should be further investigated in clinically-diagnosed samples in order to ascertain to what extent this criterion is a useful indicator of disordered gaming or if it may be best operationalized as a motivational factor for playing videogames.

The IRT analysis further indicated that criterion 6 (i.e., "*continuation*"; IGDS9-SF item 6) also presented poor fit in the sample recruited. It could be hypothesized that this criterion may be a better indicator of disordered gaming at less severe levels, which implies assigning less weight to this criterion in comparison to other diagnostic criteria. This assumption is congruent with previous studies recently conducted. For example, the study by Király et al. (2017) in a sample of 4887 Hungarian gamers found that the criteria "*continuation*", "*preoccupation*", "*negative consequences*", and "*escape*" were mostly associated with lower severity of IGD, while "*tolerance*", "*loss of control*", "*give up other activities*", and "*deception*" were mostly associated with more severe levels of IGD. Although the criteria outlined exhibited poor fit, they were still good discriminators with ordered levels, indicating that the criteria are somehow useful but with potential to measure better IGD.

Although the present findings are not definite given constraints within the design and sample of the study, the study presents preliminary evidence suggesting that not all nine criteria should have the same diagnostic weighting. This particular finding is supported by emerging evidence showing discrepancies amongst the nine IGD criteria in terms of how they contribute towards final IGD diagnosis. The results reported here could be used to inform the next revision of the nine IGD criteria in the DSM-5 in order to refine and enhance its diagnostic features. Although the results obtained were robust, there may be a few potential limitations that should be taken into consideration when interpreting the findings reported. Even though the sample was relatively large and heterogeneous, participants were recruited using a non-probability sampling technique, which potentially limits the external validity of the findings reported. Another potential limitation worth mentioning is related to the fact that in the present study no information regarding gamers' preference about the video games they played and/or video game genre was collected. By collecting such information, researchers can model their data using alternative approaches (e.g., modelling with nested data) that may be fruitful. Additionally, it is known that assessment tools of psychiatric disorders with low prevalence rates tend to present with low positive predictive values, which implies that only a small proportion of those who test positive are truly disordered (Maraz, Király, & Demetrovics, 2015). Consequently, disregarding the low positive predictive values, could potentially lead to inflated prevalence rates of a disorder, leading to unnecessary over-pathologization of behaviors. The present study also had a sample with a high percentage of male gamers, this should also be taken into account, as these findings may not be entirely representative to both genders.

Irrespective of these potential limitations, the results obtained support the validity and reliability of the Polish IGDS9-SF and its overall suitability to assess IGD in Polish speaking samples. With regards to the results obtained in the IRT analysis, it was found that some criteria might perform better than other under certain conditions. The findings of the present study will hopefully pave the way to future clinical studies aimed at ascertaining the validity of each IGD criteria and their suitability to measure this construct at different levels.

Appendix 1. Polish Internet Gaming Disorder Scale–Short-Form (IGDS9-SF)

	Nigdy	Rzadko	Czasami	Często	Bardzo często
1. Czy czujesz się zatroskany/a swoim zachowaniem związanym z grami? (Kilka przykładów: Czy rozmyślasz o swojej wcześniejszej aktywności związanej z grami? Czy uważasz, że granie stało się dominującą aktywnością w Twoim życiu codziennym?)	○	○	○	○	○
2. Czy czujesz większą drażliwość, niepokój, a nawet smutek podczas prób ograniczenia lub przerwania gry?	○	○	○	○	○
3. Czy czujesz potrzebę spędzania coraz większej ilości czasu grając, w celu osiągnięcia satysfakcji lub przyjemności?	○	○	○	○	○
4. Czy podejmowane próby ograniczenia lub zaprzestania aktywności związanej z grami kończą się niepowodzeniem?	○	○	○	○	○
5. Czy straciłeś/aś zainteresowanie poprzednim hobby lub innego typu rozrywką na skutek Twojego zaangażowania w gry?	○	○	○	○	○
6. Czy kontynuujesz granie pomimo wiedzy, że jest ono przyczyną Twoich problemów/ nieporozumień z innymi ludźmi?	○	○	○	○	○
7. Czy kiedykolwiek oszukałeś/aś, okłamałeś/aś kogokolwiek (rodzina, znajomi, erapeuci, nauczyciele itp.) z powodu Twojej aktywności związanej z graniem?	○	○	○	○	○
8. Czy grasz po to by choć na krótki czas uciec od złego humoru lub by złagodzić negatywny nastrój bezradności, poczucia winy, niepokoju itp.?	○	○	○	○	○

9. Czy przez swoją aktywność związaną z grami naraziłeś/aś się na istotne kłopoty w związku, w szkole/uczelni, w pracy?	○	○	○	○	○

Instrukcje: Poniższe pytania dotyczą Twojej aktywności jako gracza podczas ostatniego roku (tzn. w ciągu ostatnich 12 miesięcy). Aktywność gracza oznacza każdą aktywność związaną z grą odtwarzaną z komputera / laptopa, konsoli do gier lub innego urządzenia (np. telefonu komórkowego, tabletu itp.), zarówno online, jak i offline.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

APA (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

Arcelus, J., Bouman, W. P., Jones, B. A., Richards, C., Jimenez-Murcia, S., & Griffiths, M. D. (2017). Video gaming and gaming addiction in transgender people: An exploratory study. *Journal of Behavioral Addictions, 6*(1), 21–29. http://dx.doi.org/10.1556/2006.6.2017.002.

Ayesrst, L. E., & Bagby, R. M. (2011). Evaluating the psychometric properties of psychological measures. In M. M. Antony, & D. H. Barlow (Eds.). *Handbook of assessment and treatment planning of psychological disorder* (pp. 23–61). (second edition). London: The Guildford Press.

Ballabio, M., Griffiths, M. D., Urbán, R., Quartiroli, A., Demetrovics, Z., & Király, O. (2017). Do gaming motives mediate between psychiatric symptoms and problematic gaming? An empirical survey study. *Addiction Research and Theory, 25*(5), 397–408. http://dx.doi.org/10.1080/16066359.2017.1305360.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. http://dx.doi.org/10.1097/00007632-200012150-00014.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238.

Bentler, P. M., & Bonnet, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. http://dx.doi.org/10.1037/0033-2909.88.3.588.

Bijvank, M. N., Konijn, E. A., & Bushman, B. J. (2012). "We don't need no education": Video game preferences, video game motivations, and aggressiveness among adolescent boys of different educational ability levels. *Journal of Adolescence, 35*(1), 153–162. http://dx.doi.org/10.1016/j.adolescence.2011.04.001.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*(4), 443–459. http://dx.doi.org/10.1007/bf02293801.

Brunborg, G. S., Mentzoni, R. A., & Frøyland, L. R. (2014). Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems? *Journal of Behavioral Addictions, 3*(1), 27–32. http://dx.doi.org/10.1556/JBA.3.2014.002.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York: Springer-Verlag.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*(6), 1–29. http://dx.doi.org/10.18637/jss.v048.i06.

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement, 41*(3), 178–194. http://dx.doi.org/10.1177/0146621616677520.

Corp, I. B. M. (2015). *IBM SPSS statistics for windows, version 23.* New York: IBM Corp.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* New York: Lawrence Erlbaum Associates.

Festl, R., Scharkow, M., & Quandt, T. (2013). Problematic computer game use among adolescents, younger and older adults. *Addiction, 108*(3), 592–599. http://dx.doi.org/10.1111/add.12016.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (fourth edition). London: SAGE Publications Ltd.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. http://dx.doi.org/10.2307/3151312.

Frölich, J., Lehmkuhl, G., Orawa, H., Bromba, M., Wolf, K., & Görtz-Dorten, A. (2016). Computer game misuse and addiction of adolescents in a clinically referred study sample. *Computers in Human Behavior, 55*(Part A), 9–15. http://dx.doi.org/10.1016/j.chb.2015.08.043.

Fuster, H., Carbonell, X., Pontes, H. M., & Griffiths, M. D. (2016). Spanish validation of the Internet Gaming Disorder-20 (IGD-20) test. *Computers in Human Behavior, 56*, 215–224. http://dx.doi.org/10.1016/j.chb.2015.11.050.

Gomez, R., Stavropoulos, V., Beard, C., & Pontes, H. M. (2018). Item response theory analysis of the recoded Internet Gaming Disorder scale-short-form (IGDS9-SF). *International Journal of Mental Health and Addiction*. http://dx.doi.org/10.1007/s11469-018-9890-z.

Griffiths, M. D., King, D. L., & Demetrovics, Z. (2014). DSM-5 Internet Gaming Disorder needs a unified approach to assessment. *Neuropsychiatry, 4*(1), 1–4. http://dx.doi.org/10.2217/npy.13.82.

Hagström, D., & Kaldo, V. (2014). Escapism among players of MMORPGs—Conceptual clarification, its relation to mental health factors, and development of a new measure. *Cyberpsychology, Behavior and Social Networking, 17*(1), 19–25. http://dx.doi.org/10.1089/cyber.2012.0222.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis. A global perspective* (seventh edition). Upper Saddle River, CA: Pearson Prentice Hall.

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Education Measurement: Issues and Practice, 12*(3), 38–47. http://dx.doi.org/10.1111/j.1745-3992.1993.tb00543.x.

Hawi, N. S., & Samaha, M. (2017). Validation of the Arabic version of the Internet Gaming Disorder-20 test. *Cyberpsychology, Behavior and Social Networking, 20*(4), 268–272. http://dx.doi.org/10.1089/cyber.2016.0493.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. http://dx.doi.org/10.1080/10705519909540118.

Kahn, A. S., Shen, C., Lu, L., Ratan, R. A., Coary, S., Hou, J., et al. (2015). The Trojan player typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations. *Computers in Human Behavior, 49*, 354–361. http://dx.doi.org/10.1016/j.chb.2015.03.018.

Kang, T., & Chen, T. T. (2008). Performance of the generalized S-X2 item fit index for polytomous IRT models. *Journal of Educational Measurement, 45*(4), 391–406. http://dx.doi.org/10.1111/j.1745-3984.2008.00071.x.

Kaptsis, D., King, D. L., Delfabbro, P. H., & Gradisar, M. (2016a). Trajectories of abstinence-induced internet gaming withdrawal symptoms: A prospective pilot study. *Addictive Behaviors Reports, 4*, 24–30. http://dx.doi.org/10.1016/j.abrep.2016.06.002.

Kaptsis, D., King, D. L., Delfabbro, P. H., & Gradisar, M. (2016b). Withdrawal symptoms in Internet Gaming Disorder: A systematic review. *Clinical Psychology Review, 43*, 58–66. http://dx.doi.org/10.1016/j.cpr.2015.11.006.

King, D. L., & Delfabbro, P. H. (2016). Defining tolerance in Internet Gaming Disorder: Isn't it time? *Addiction, 111*(11), 2064–2065. http://dx.doi.org/10.1111/add.13448.

Király, O., Sleczka, P., Pontes, H. M., Urbán, R., Griffiths, M. D., & Demetrovics, Z. (2017). Validation of the ten-item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 Internet Gaming Disorder criteria. *Addictive Behaviors, 64*, 253–260. http://dx.doi.org/10.1016/j.addbeh.2015.11.005.

Király, O., Urbán, R., Griffiths, M. D., Ágoston, C., Nagygyörgy, K., Kökönyei, G., et al. (2015). The mediating effect of gaming motivation between psychiatric symptoms and problematic online gaming: An online survey. *Journal of Medical Internet Research, 17*(4), e88. http://dx.doi.org/10.2196/jmir.3515 (Original paper).

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (third edition). New York, NY: The Guildford Press.

Ko, C. H., Yen, J. Y., Chen, S. H., Wang, P. W., Chen, C. S., & Yen, C. F. (2014). Evaluation of the diagnostic criteria of Internet Gaming Disorder in the DSM-5 among young adults in Taiwan. *Journal of Psychiatric Research, 53*(6), 103–110. http://dx.doi.org/10.1016/j.jpsychires.2014.02.008.

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017a). Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: Issues, concerns, and recommendations for clarity in the field. *Journal of Behavioral Addictions, 6*(2), 103–109. http://dx.doi.org/10.1556/2006.5.2016.062.

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017b). DSM-5 diagnosis of Internet Gaming Disorder: Some ways forward in overcoming issues and concerns in the gaming studies field. *Journal of Behavioral Addictions, 6*(2), 133–141. http://dx.doi.org/10.1556/2006.6.2017.032.

Leménager, T., Dieter, J., Hill, H., Hoffmann, S., Reinhard, I., Beutel, M., et al. (2016). Exploring the neural basis of avatar identification in pathological Internet gamers and of self-reflection in pathological social network users. *Journal of Behavioral Addictions,* 1–15. http://dx.doi.org/10.1556/2006.5.2016.048.

Lemmens, J. S., Valkenburg, P. M., & Gentile, D. A. (2015). The Internet Gaming Disorder Scale. *Psychological Assessment, 27*(2), 567–582. http://dx.doi.org/10.1037/pas0000062.

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2009). Development and validation of a game addiction scale for adolescents. *Media Psychology, 12*(1), 77–95. http://dx.doi.org/10.1080/15213260802669458.

Li, W., Garland, E. L., McGovern, P., O'Brien, J. E., Tronnier, C., & Howard, M. O. (2017). Mindfulness-oriented recovery enhancement for Internet Gaming Disorder in U.S. adults: A stage I randomized controlled trial. *Psychology of Addictive Behaviors, 31*(4), 393–402. http://dx.doi.org/10.1037/adb0000269.

Lim, J.-A., Lee, J.-Y., Jung, H. Y., Sohn, B. K., Choi, S.-W., Kim, Y. J., et al. (2016). Changes of quality of life and cognitive function in individuals with Internet Gaming Disorder: A 6-month follow-up. *Medicine, 95*(50), e5695. http://dx.doi.org/10.1097/md.0000000000005695.

Männikkö, N., Billieux, J., Nordström, T., Koivisto, K., & Kääriäinen, M. (2017). Problematic gaming behaviour in Finnish adolescents and young adults: Relation to game genres, gaming motives and self-awareness of problematic use. *International Journal of Mental Health and Addiction, 15*(2), 324–338. http://dx.doi.org/10.1007/

s11469-016-9726-7.
Maraz, A., Király, O., & Demetrovics, Z. (2015). Commentary on: Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research the diagnostic pitfalls of surveys: If you score positive on a test of addiction, you still have a good chance not to be addicted. *Journal of Behavioral Addictions, 4*(3), 151–154. http://dx.doi.org/10.1556/2006.4.2015.026.
Monacis, L., De Palo, V., Griffiths, M. D., & Sinatra, M. (2016). Validation of the Internet Gaming Disorder Scale – Short-Form (IGDS9-SF) in an Italian-speaking sample. *Journal of Behavioral Addictions, 5*(4), 683–690. http://dx.doi.org/10.1556/2006.5.2016.083.
Monacis, L., de Palo, V., Griffiths, M. D., & Sinatra, M. (2017). Social networking addiction, attachment style, and validation of the Italian version of the Bergen Social Media Addiction Scale. *Journal of Behavioral Addictions,* 1–9. http://dx.doi.org/10.1556/2006.6.2017.023.
Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *ETS Research Report Series, 1992*(1), http://dx.doi.org/10.1002/j.2333-8504.1992.tb01436.x.
Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (seventh edition). Los Angeles, CA: Muthén & Muthén.
Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement, 24*(1), 50–64. http://dx.doi.org/10.1177/01466216000241003.
Pápay, O., Urbán, R., Griffiths, M. D., Nagygyörgy, K., Farkas, J., Kökönyei, G., et al. (2013). Psychometric properties of the problematic online gaming questionnaire short-form and prevalence of problematic online gaming in a national sample of adolescents. *Cyberpsychology, Behavior and Social Networking, 16*(5), 340–348. http://dx.doi.org/10.1089/cyber.2012.0484.
Park, S., Jeon, H. J., Son, J. W., Kim, H., & Hong, J. P. (2017). Correlates, comorbidities, and suicidal tendencies of problematic game use in a national wide sample of Korean adults. [journal article]. *International Journal of Mental Health Systems, 11*(1), 35. http://dx.doi.org/10.1186/s13033-017-0143-5.
Petry, N. M., & O'Brien, C. P. (2013). Internet Gaming Disorder and the DSM-5. *Addiction, 108*(7), 1186–1187. http://dx.doi.org/10.1111/add.12162.
Petry, N. M., Rehbein, F., Ko, C. H., & O'Brien, C. P. (2015). Internet Gaming Disorder in the DSM-5. *Current Psychiatry Reports, 17*(72), 1–9. http://dx.doi.org/10.1007/s11920-015-0610-0.
Pontes, H. M., & Griffiths, M. D. (2014). Assessment of Internet Gaming Disorder in clinical research: Past and present perspectives. *Clinical Research and Regulatory Affairs, 31*(2–4), 35–48. http://dx.doi.org/10.3109/10601333.2014.962748.
Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 Internet Gaming Disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior, 45*, 137–143. http://dx.doi.org/10.1016/j.chb.2014.12.006.
Pontes, H. M., & Griffiths, M. D. (2016). Portuguese validation of the Internet Gaming Disorder Scale–Short-Form. *CyberPsychology, Behavior & Social Networking, 19*(4), 288–293. http://dx.doi.org/10.1089/cyber.2015.0605.
Pontes, H. M., Király, O., Demetrovics, Z., & Griffiths, M. D. (2014). The conceptualisation and measurement of DSM-5 Internet Gaming Disorder: The development of the IGD-20 test. *PLoS One, 9*(10), e110137. http://dx.doi.org/10.1371/journal.pone.0110137.
Pontes, H. M., Kuss, D. J., & Griffiths, M. D. (2015). Clinical psychology of internet addiction: A review of its conceptualization, prevalence, neuronal processes, and implications for treatment. *Neuroscience and Neuroeconomics, 4*, 11–23. http://dx.doi.org/10.2147/NAN.S60982.
Pontes, H. M., Kuss, D. J., & Griffiths, M. D. (2017). Psychometric assessment of Internet Gaming Disorder in neuroimaging studies: A systematic review. In C. Montag, & M. Reuter (Eds.). *Internet addiction: Neuroscientific approaches and therapeutical implications including smartphone addiction* (pp. 181–208). Cham: Springer International Publishing.
Pontes, H. M., Macur, M., & Griffiths, M. D. (2016a). Construct validity and preliminary psychometric properties of the Internet Gaming Disorder Scale – Short-Form (IGDS9-SF) among Slovenian youth: A nationally representative study. *Journal of Behavioral Addictions, 5*(s1), 35. http://dx.doi.org/10.1556/JBA.5.2015.Suppl.1.
Pontes, H. M., Macur, M., & Griffiths, M. D. (2016b). Internet Gaming Disorder among Slovenian primary schoolchildren: Findings from a nationally representative sample of adolescents. *Journal of Behavioral Addictions, 5*(2), 304–310. http://dx.doi.org/10.1556/2006.5.2016.042.
Rehbein, F., Kliem, S., Baier, D., Mößle, T., & Petry, N. M. (2015). Prevalence of Internet Gaming Disorder in German adolescents: Diagnostic contribution of the nine DSM-5 criteria in a state-wide representative sample. *Addiction, 110*(5), 842–851. http://dx.doi.org/10.1111/add.12849.
Rubin, A., & Babbie, E. (2009). *Essential research methods for social work* (second edition). Belmont, CA: Cengage Learning.
Samejima, F. (1997). Graded Response Model. In W. J. van der Linden, & R. K. Hambleton (Eds.). *Handbook of Modern Item Response Theory*. New York, NY: Springer.
Schneider, L. A., King, D. L., & Delfabbro, P. H. (2017). Maladaptive coping styles in adolescents with Internet Gaming Disorder symptoms. *International Journal of Mental Health and Addiction,* 1–12. http://dx.doi.org/10.1007/s11469-017-9756-9.
Sibley, C. G. (2012). The Mini-IPIP6: Item response theory analysis of a short measure of the big-six factors of personality in New Zealand. *New Zealand Journal of Psychology, 41*(3), 21–31.
Sioni, S. R., Burleson, M. H., & Bekerian, D. A. (2017). Internet Gaming Disorder: Social phobia and identifying with your virtual self. *Computers in Human Behavior, 71*, 11–15. http://dx.doi.org/10.1016/j.chb.2017.01.044.
Starcevic, V. (2017). Internet Gaming Disorder: Inadequate diagnostic criteria wrapped in a constraining conceptual model. *Journal of Behavioral Addictions, 6*(2), 110–113. http://dx.doi.org/10.1556/2006.6.2017.012.
Stavropoulos, V., Beard, C., Griffiths, M. D., Buleigh, T., Gomez, R., & Pontes, H. M. (2017). Measurement invariance of the Internet Gaming Disorder Scale–Short-Form (IGDS9-SF) between Australia, the USA, and the UK. *International Journal of Mental Health and Addiction*. http://dx.doi.org/10.1007/s11469-017-9786-3.
Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (sixth edition). Essex, England: Pearson Education.
Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. http://dx.doi.org/10.18637/jss.v045.i03.
Wartberg, L., Kriston, L., & Kammerl, R. (2017). Associations of social support, friends only known through the Internet, and health-related quality of life with Internet Gaming Disorder in adolescence. *Cyberpsychology, Behavior and Social Networking, 20*(7), 436–441. http://dx.doi.org/10.1089/cyber.2016.0535.
Wartberg, L., Kriston, L., Kramer, M., Schwedler, A., Lincoln, T., & Kammerl, R. (2017). Internet Gaming Disorder in early adolescence: Associations with parental and adolescent mental health. *European Psychiatry, 43*, 14–18. http://dx.doi.org/10.1016/j.eurpsy.2016.12.013.
Wartberg, L., Kriston, L., & Thomasius, R. (2017). The prevalence and psychosocial correlates of Internet Gaming Disorder - analysis in a nationally representative sample of 12- to 25-year-olds. *Deutsches Ärzteblatt International, 114*(25), 419–424. http://dx.doi.org/10.3238/arztebl.2017.0419.
Wittek, C. T., Finserås, T. R., Pallesen, S., Mentzoni, R. A., Hanss, D., Griffiths, M. D., et al. (2016). Prevalence and predictors of video game addiction: A study based on a national representative sample of gamers. *International Journal of Mental Health and Addiction, 14*(5), 672–686. http://dx.doi.org/10.1007/s11469-015-9592-8.
Wu, A. M. S., Lai, M. H. C., Yu, S., Lau, J. T. F., & Lei, M.-W. (2017). Motives for online gaming questionnaire: Its psychometric properties and correlation with Internet Gaming Disorder symptoms among Chinese people. *Journal of Behavioral Addictions, 6*(1), 11–20. http://dx.doi.org/10.1556/2006.6.2017.007.
Wu, T. Y., Lin, C.-Y., Årestedt, K., Griffiths, M. D., Broström, A., & Pakpour, A. H. (2017). Psychometric validation of the Persian nine-item Internet Gaming Disorder Scale – Short Form: Does gender and hours spent online gaming affect the interpretations of item descriptions? *Journal of Behavioral Addictions, 6*(2), 256–263. http://dx.doi.org/10.1556/2006.6.2017.025.
Yee, N. (2006). Motivations for play in online games. *Cyberpsychology & Behavior, 9*(6), 772–775. http://dx.doi.org/10.1089/cpb.2006.9.772.
Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8*(2), 125–145. http://dx.doi.org/10.1177/014662168400800201.
Yu, H., & Cho, J. (2016). Prevalence of Internet Gaming Disorder among Korean adolescents and associations with non-psycothic psychological symptoms, and physical aggression. *American Journal of Health Behavior, 40*(6), 705–716. http://dx.doi.org/10.5993/AJHB.40.6.3.